

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11)

EP 0 969 014 A1

(12)

EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:

05.01.2000 Bulletin 2000/01

(51) Int. Cl.⁷: C07K 14/405, C12N 9/10,

C12N 15/54, C12N 15/31,

C12N 1/21

// (C12N15/54, C12R1:89),

(C12N9/10, C12R1:19)

(21) Application number: 98900438.7

(22) Date of filing: 20.01.1998

(86) International application number:

PCT/JP98/00194

(87) International publication number:

WO 98/32770 (30.07.1998 Gazette 1998/30)

(84) Designated Contracting States:
BE CH DE DK ES GB IT LI NL

(72) Inventors:

• FERRI, R., Stefano
I-37121 Verona (IT)

• TOGURI, Toshihiro; Kirin Beer Kabushiki Kaisha
Yokohama-shi, Kanagawa 236 (JP)

(30) Priority: 24.01.1997 JP 1143097

(74) Representative: HOFFMANN - EITLE
Patent- und Rechtsanwälte
Arabellastrasse 4
81925 München (DE)

(71) Applicant:
KIRIN BEER KABUSHIKI KAISHA
Chuo-Ku, Tokyo 104 (JP)

(54) Sg(b)-KETOACYL-ACP-SYNTETASE II ENZYME AND GENE ENCODING THE SAME

(57) A gene of a protein having an enzymatic activity enabling the control or regulation of the contents of saturated and unsaturated fatty acids in plant cells and the enzyme protein as the expression product of this gene, more specifically: proteins showing a β -ketoacyl-ACP-synthetase II (KAS II) activity and having a specific amino acid sequence typified by the amino acid

sequence of KAS II originating in blue-green algae (*Anacystis nidulans*) or an amino acid sequence substantially equivalent thereto; a KAS II protein gene encoding the amino acid sequence of the above protein; and recombinant vectors containing this gene and cells having this gene transferred thereinto.

MTETGGRMV ITGLGAI TP1 GNDPTEYHOG ILAGRNGIDL IRGPFDASRHA CKIAGEVKDF
DPTQYMDRKD AKMMDRFAQL AVAASRQAVYA DAKLDITELN ADAIGVLLGS GIGGLRMED
OATVILLEKGP DRCSPPRIVPMI MIAAMAAAGLT A1OLGAKGPC MVTYTACAAG SNAVGAEFRL
IOMGYADAMI CGGTESCVPF LAMAGFAACK ALSLRNDOPA HACRPFDODR DGFVWEGAG
ILVLESLEHA QARGAHLYGE IVGYGMTCDA YH1TSPVPGG LGAAATAEFG LRDANLPSO
VSYINAHGTS TPANDSTETA AIKKALGEHA YKTVISSTKS WTHGLLGGSG GIEAVAATLA
IAEDWMPPTI NLEDPOPDCD LDYVPMQAPS LPYEVALSRS FGFGGHNVTI AFRKRP

F I G. 1

P 0 969 014 A1

DescriptionTechnical Field

5 [0001] The present invention relates to the amino acid sequence of a synthetase responsible for the synthesis of fatty acids in plants and the structure of DNA related to the same, that is to say, a fatty acid synthesizing enzyme protein having a specific amino acid sequence, and a gene coding for it. Cells can be transformed with such a gene and chimeric genes in which an appropriate regulatory sequence (regulatory gene) has been inserted to control the amounts of saturated and unsaturated fatty acids in the cell.

Background Art

[0002] Fatty acid synthases are known to be divided into two types; the enzymes in animals and yeasts are fatty acid synthetase complexes (FAS) in which a variety of enzymes are wholly linked as a complex having a single function (type I), while those in higher plant cells and prokaryotes are of such type that each of the enzymes become independently disconnected outside the organisms (type II). An acyl carrier protein (ACP) which is a soluble protein is required for the synthesis of a fatty acid with the enzyme type II, and the fatty acids are synthesized as an acyl-ACP. The final product of the synthesis system is palmitoyl-ACP. The palmitoyl-ACP is further converted into stearoyl-ACP by chain elongation before desaturation with a soluble fatty acid desaturase (stearoyl-ACP desaturase) to lead to oleoyl-ACP. Palmitic acid and oleic acid are incorporated into polar lipid, and then the latter is further desaturated (J. Ohlrogge and J. Browse (1995) Lipid Biosynthesis. The Plant Cell, 7, p. 957-970).

[0003] The chain elongation enzyme which catalyzes the chain elongation from palmitoyl-ACP to stearoyl-ACP produces stearoyl-ACP from palmitoyl-ACP, malonyl-ACP and NADPH (J. Ohlrogge and J. Browse (1995) Lipid Biosynthesis. The Plant Cell, 7, p. 957-970). These reactions describe the total scheme of a series of enzyme reactions for the production of a stearoyl-thioester by the reduction, dehydration and further reduction of the condensation product of a palmitoyl-thioester and the C2 unit.

[0004] Lipid biosynthesis in plants has been studied very extensively (Browse et al., Annu. Rev. Plant Physiol. Mol. Biol. (1991) 42: 467-506). It has been elucidated from these researches that the production of stearic acid starting from palmitic acid in plant cells is a reaction catalyzed by the enzyme β -ketoacyl-ACP synthetase II (KASII). However, there has been described in the aforementioned publication the isolation of neither the enzyme KASII or its gene, and thus their sequences remain unknown.

[0005] Three isozymes of KAS have been found in chloroplast in plants. Among the two isozymes other than KASII, the isozyme KASIII catalyzes the initiation of the synthesis of an acyl chain, while the isozyme KASI catalyzes the elongation reaction of an acyl chain to the palmitoyl-ACP with 16 carbon atoms. A variety of mutants of enzymes involved in the lipid synthesis of plants have been isolated from *Arabidopsis*, among which the enzymes responsible for the desaturation reaction have been studied extensively. There have also been described for the KASII, the mutant of which has been designated as fab1. In this mutant, the KASII enzyme activity was lowered to 65%, and thus the palmitic acid content increased by 7% in leaves and 3% in roots (Wu et al., Plant Physiol. (1994) 106: 143-150). As regards the complete purification of the enzyme KASII, genes have been cloned from a castor bean (*Ricinus communis*; Japanese Patent 501446/1995) and soybean (*Glycine max*; Japanese Patent Laid-Open Publication No. 500234/1994) on the basis of the amino acid sequences of the limitedly degraded peptide of the purified enzyme. In the above described publications, as regards the changes of fatty acids by transformation with these genes, the C16 fatty acid content on the expression of the gene of the castor bean in *E. coli* was decreased by ca. 20% thus corresponding to a little over 30% of the total fatty acid content, while on introducing the soybean gene into canola the palmitic acid content in the seeds was decreased by 0.8% and on introducing the gene into tobacco the palmitic acid content in the leaves was decreased by ca. 2%. In this connection, no sequence exhibiting distinct homology has been curiously found between the genes of the castor bean and soybean.

[0006] By the way, it has been known that in membrane lipids constituting biomembrane, the phase transition temperature varies primarily depending on the unsaturation degrees of the fatty acid linked to the lipid, and as a result the chilling resistance of the organism also varies. It is thought that the unsaturation degree of the membrane lipid is effectively increased with an enzyme such as fatty acid acyltransferase (PCT/JP 92/00024 (PCT/WO 92/13082)), fatty acid desaturase (PCT/JP 94/02288 (PCT/WO 95/18222)).

Disclosure of the Invention

[0007] In consideration of the above described situations, the object of the present invention is to provide a gene of a protein having an enzyme activity which makes it possible to regulate or control the content of saturated fatty acids and unsaturated fatty acids in plant cells or microorganism cells and an enzyme protein as the expression product.

[0008] It is believed that if there is a protein having such an enzyme activity that the decrease of the enzyme activity responsible for the synthesis of fatty acids leads to the increase of the palmitic acid content in lipids of cells while the increase of enzyme activity leads to the increase of content of the fatty acids with 18 or more carbon atoms, the unsaturated fatty acid contents in the lipids is possibly increased for example as a result of the increase of content of the fatty acids with 18 or more carbon atoms due to the increase of the enzyme activity.

[0009] The present inventors have conducted earnest researches in order to solve the above described problems, and as a result, successfully isolated a gene which codes for an enzyme β -ketoacyl-ACP synthetase II (KASII) from cyanobacterium (Anacystis nidulans), and found that the introduction of the gene into E. coli confers the KASII producing ability whereby fatty acids having extended chain length increase. The present invention has been accomplished on the basis of the finding.

[0010] That is, the present invention relates to the protein which has an amino acid sequence represented by SEQ ID NO. 2 or substantially the same amino acid sequence as the one represented by SEQ ID NO. 2 and exhibit the KASII enzyme activity.

[0011] The present invention also relates to the KASII enzyme gene coding for the protein which has an amino acid sequence represented by SEQ ID NO. 2 or substantially the same amino acid sequence as the one represented by SEQ ID NO. 2 and exhibit the KASII enzyme activity.

[0012] Furthermore, the present invention relates to the recombinant vector containing the gene and the cells in which the gene has been introduced.

20 Brief Description of the Drawings

[0013]

Fig. 1 shows the amino acid sequence represented by single-letter codes corresponding to SEQ ID NO. 2 (three-letter codes amino acid).

25 Fig. 2 illustrates the comparative chart of the amino acid sequence of barley (Accession No. P23902 in SWISS-PROT data base) KASI enzyme with the amino acid sequence expected from the DNA sequence of the castor bean (GENBANK Accession No. L13241) KASII enzyme.

30 Fig. 3 illustrates the comparative chart of the amino acid sequence of the enzyme KASII (referred to as KASII in the drawing) derived from Anacystis nidulans in the present invention with the amino acid sequence (referred to as Fab) of fabF or J (code S11 1069) of Synechocystis sp. Strain PCC6803.

Best Mode for carrying out the Invention

35 [0014] The present invention is now explained in detail as follows.

KASII enzyme active protein and its gene

40 [0015] As described above, the protein of the present invention having the KASII enzyme activity has the amino acid sequence represented by SEQ ID NO. 2 (corresponding to the amino acid sequence (single-letter code) in Fig. 1) or substantially the same amino acid sequence as the one represented by SEQ ID NO. 2, and the gene of the KASII enzyme active protein according to the present invention codes for the above described protein having the amino acid sequence represented by SEQ ID NO. 2 or substantially the same amino acid sequence as the one represented by SEQ ID NO. 2. The phraseology "protein having the KASII enzyme activity" or "KASII enzyme active protein" in the present invention means the protein having an enzyme activity which produces a longer fatty acid (particularly stearic acid) by extending the chain length of a fatty acid (particularly palmitic acid).

45 [0016] As the KASII enzyme active proteins described above, there can be used in the present invention an appropriate naturally occurring gene products as well as mutant gene products in which a part of the amino acid sequences of the proteins has been mutated, provided that the proteins have the above described KASII enzyme activity. By way of example, the product of the KASII enzyme active protein gene typically includes the KASII enzyme derived from the cyanobacterium as a microorganism (SEQ ID NO. 2).

50 [0017] The phrase "substantially the same amino acid sequence" in the present invention means that the sequence of the mutant described above is also included, such as typically the amino acid sequence of the enzyme protein derived from the cyanobacterium represented by SEQ ID NO. 2 (SEQ ID NO. 2, Fig. 1) or the sequence in which one or more, preferably one or a few amino acids have been substituted, deleted, inserted or added.

55 [0018] Therefore, the term "substantially" in the case of "the gene coding for ... substantially the same amino acid sequence" in the present invention is intended to include not only the gene of a DNA sequence coding for the naturally occurring protein having the KASII enzyme activity defined above, but also the gene of a DNA sequence coding for the

mutant KASII enzyme active protein described above, typically the gene of the DNA sequence coding for the amino acid sequence represented by SEQ ID NO. 2 or the amino acid sequence in which one or more, preferably one or a few amino acids are substituted, deleted, inserted or added. Also, it goes without saying that when a DNA chain generally codes for a polypeptide having an amino acid sequence, plural gene codes (codons) corresponding to an amino acid are present (degenerated mutants), and thus any gene codes can be used also in the DNA chain coding for the KASII enzyme active protein of the present invention.

[0019] The KASII enzyme active protein encoded by the gene of the present invention has a function of the chain elongation enzyme for the fatty acid synthesis which is originally present in plants and microorganisms as described above, and thus, in the further specific function, has an enzyme activity for producing a longer fatty acid (particularly, stearic acid) that the chain length of a fatty acid (particularly, palmitic acid) has been elongated. The typical example of the protein according to the present invention is the one derived from cyanobacterium. The chemical structure of the enzyme KASII derived from the cyanobacterium is locally similar to the protein encoded by the KASI gene of *E. coli* and barley, and also similar to the enzyme derived from the castor bean among the patent publications regarding to the above described KASII. The KASII enzyme active protein according to the present invention has, as described above, the amino acid sequence represented by SEQ ID NO. 2 or substantially the same amino acid sequence as the above described one, and a markedly high enzyme activity for producing a longer fatty acid (particularly, stearic acid) in which the chain length of a fatty acid (particularly, palmitic acid) has been elongated.

[0020] While a means for obtaining the gene coding for the protein of the present invention is the chemical synthesis of at least a part of the chain according to the method of nucleic acid synthesis, it is more preferable in consideration of the number of the amino acids to be linked to use the method, rather than the chemical synthesis, in which cDNA is synthesized from mRNA isolated from a naturally occurring material, in particular cyanobacterium as a bacterium, and the gene is obtained from the gene library by the method commonly used in the field of genetic engineering.

[0021] The gene of the enzyme KASII can be obtained for example as follows.

[0022] Firstly, the enzyme from a higher plant or a microorganism, particularly cyanobacterium is purified by the known method, and fragmented with peptidase to determine the amino acid sequences of the fragments. Oligonucleotides corresponding to the fragmented peptides whose amino acid sequence has been determined are then synthesized. The total RNA is separately extracted from the plant or the microorganism, and the DNA complementary to the RNA (cDNA) is synthesized. The cDNA is linked to an appropriate vector such as phage λ gt11 to make the cDNA library. In this connection, as the method for screening the gene, the conventional methods, for example immunological methods such as the plaque hybridization method with antibody or the colony hybridization method, or the hybridization method with nucleotide probe or the like can be used.

[0023] It is also possible to obtain the target sequence by designing primers corresponding to the short DNA sequences positioned at both of the ends of the aimed sequence on the basis of the consensus sequence of a known KASII enzyme or of the other isozyme related to it, and conducting PCR with DNA obtained from a material used for determining the total sequence as a template. In this case, the activity can be identified for example by expressing the KAS gene in *E. coli* in order to discriminating the isozyme of the gene.

[0024] The DNA sequence of the gene according to the present invention in the clone thus screened can be determined and confirmed generally by the known methods such as the dideoxynucleotide chain terminating method with M13 phage (Sambrook et al., Molecular Cloning, 2nd edition (1989)).

[0025] The present gene of which DNA sequence has been determined as described above can be also synthesized generally by the known means, for example a commercially available DNA synthesizer by the phosphite method.

[0026] Also, for expressing a DNA chain or its fragment to produce a protein or a polypeptide encoded thereby, an expression regulatory sequence is required in addition to the DNA sequence (coding region) corresponding to the amino acid sequence. Thus, the DNA chain of the present invention includes the DNA sequence comprising such expression regulatory sequence. Among the expression regulatory region, an important one particularly for expressing it in a higher plant is the promoter sequence upstream of the coding region (e.g. derived from the 35S promoter of cauliflower mosaic virus), and the poly A addition signal downstream (e.g. derived from the terminator of nopaline synthesis enzyme). When DNA obtained is the genome gene of a higher plant, it can also be used directly provided that the DNA sequence comprises expression regulatory region.

Use of KASII enzyme active protein gene

[0027] As described above, the present invention relates also to the recombinant vector comprising the above described DNA chain or its fragment, and to cells into which the gene has been introduced.

[0028] The recombinant vector is a vector to which the above described DNA chain or its fragment has been linked, and there can be used the known vector such as plasmid (e.g. pET17b), phage (e.g. λ ZAPII).

[0029] The enzyme KASII can be produced in a host such as an appropriate plant or microorganism cells as described above by introducing the recombinant vector DNA into the host for expression.

[0030] In this connection, although the cells may be either microorganism cells or plant cells irrespective of the kind of organisms, the microorganism cells include *E. coli* and the like, and the plant cells include chilling sensitive plants such as tobacco and the like. The gene can be introduced into plants generally with the known methods such as the ones described in "Plant Molecular Biology Manual, Second Edition; S.G. Gelvin, and R.A. Schilperoort eds, Kluwer Academic Publishers, 1995". By way of example, there can be mentioned the biological methods which include a method with virus or a method with Agrobacterium, and the physicochemical methods which include the electroporation method, the polyethylene glycol method, the particle gun method, and the like.

[0031] Also, the enzyme KASII is a protein which is present in chloroplast envelope in plants, so that it is necessary to attach the DNA chain coding for transit peptide to the chloroplast upstream the enzyme KASII. By way of example, the small sub-unit gene of ribulose-1,5-bisphosphate carboxylase of pea can be used as a gene coding for the transit peptide.

[0032] The KASII enzyme active protein of the present invention, typically the gene coding for the KASII enzyme active protein (amino acid SEQ ID NO. 2) derived from the cyanobacterium *Anacystis nidulans* (the DNA sequence of the gene derived from the cyanobacterium is represented by SEQ ID NO. 1) is useful for the improvement of lipid composition in plants and microorganisms by transformation, particularly the control of the amount ratio between fatty acids with 16 and 18 carbon atoms.

[0033] The expression of the KASII enzyme protein of the present invention as a foreign protein in an organism leads to the elongation of the chain length of fatty acids from 16 carbon atoms (palmitic acid) to the 18 carbon atoms (stearic acid), the stearic acid is desaturated in an organism, and the content of unsaturated fatty acids is increased. It is believed that chilling resistance is enhanced in plants in which unsaturated fatty acids have been increased (PCT/WO 92/13082, PCT/WO 95/18222), and it is expected that the resistance to the stress of culture is enhanced in microorganisms (yeast, *E. coli* and the like) in which unsaturated fatty acids have been increased.

Examples

[0034] The present invention is described below in detail with reference to examples, it is not to be limited by these examples.

Example 1 Preparation of DNA derived from *A. nidulans*, and preparation of DNA library

[0035] *A. nidulans* (Catalog No. IAM M-6: it is possible to obtain it from Institute of Molecular Cytology, Tokyo University) was cultured in about 100 ml of the BG-11 culture medium prepared according to the method described in p. 279 of Plant Molecular Biology, by Shaw (IRL PRESS, 1988). The bacterial cells were cultured sufficiently by shaking at 120 times/min under a fluorescent light of 1000 lux at 25°C. The cells were recovered by centrifugation at 5,000 g for 10 min. at room temperature.

[0036] In order to isolate DNA, the precipitated cells were suspended in 50 ml of 50 mM Tris Cl (pH 8.0), 1 mM EDTA (solution A) and washed by centrifugation again. The cells were then re-suspended again in 15 ml (solution) of 50 mM Tris Cl (pH 8.0), 20 mM EDTA, 50 mM NaCl, 0.25 M sucrose (Solution B), to which 40 mg of lysozyme (Sigma) dissolved in Solution B was added, and the mixture was shaken slowly at 37°C. After 1 hour, 15 mg of proteinase K and SDS at a final concentration of 1% were added, and the mixture was shaken slowly over night at 37°C. Next day, NaClO₄ was adjusted to a concentration of 1 M, 20 ml of chloroform/isoamyl alcohol (24 : 1) was added, the mixture was shaken slowly for 10 minutes, and the aqueous layer was separated by centrifugation. After extraction with chloroform/isoamyl alcohol was repeated once again, 50 ml of ethanol was added, and DNA was recovered by winding it around a glass rod. The DNA was dissolved in 20 ml of solution A, NaCl was adjusted to a concentration of 0.1 M, RNase at a concentration of 50 mg/ml was added and the reaction was conducted at 37°C for 1 hour. The reaction mixture was then subjected to an extraction twice with an equivalent amount of phenol saturated with solution A. After DNA in the aqueous layer was recovered by the addition of ethanol and washed with 70% ethanol, it was dissolved in 1 ml of solution A to prepare the DNA solution.

[0037] After partial digestion of ca. 100 g of DNA with Sau 3A I for the purpose of preparing a genomic DNA library from DNA thus obtained, DNA of about 9 - 23 kb was collected by ultracentrifugation on a gradient of sucrose density according to the method described by Sambrook et al. It was cloned into DASH II (kit by Stratagene) cleaved with Bam HI and Hind III.

Example 2 Cloning of KASII enzyme-like gene from cyanobacterium *Anacystis nidulans*

[0038] Several short DNA chains were synthesized by comparing the enzyme KASI of barley with the enzyme KASII of castor bean while paying attention to the regions having high homology between these enzymes (Fig. 2). Among these chains, distinct bands in accordance with expected sizes were observed in reactions carried out with the following

combination.

1: 5'-CC(ACGT)CC(AG)AA(ACGT)CC(AG)AA(ACGT)GA(AG)TT-3' (SEQ ID NO. 3)
2: 5'-GA(AG)GA(AG)GT(ACGT)AA(CT)TA(CT)AT(ACT)AA(CT)GC-3' (SEQ ID NO. 4)

[0039] Among the sequences, SEQ ID NO. 4 is a sense primer corresponding to the amino acid sequence EEVNYINA, and SEQ ID NO. 3 is a primer coding for the anti-sense chain corresponding to the amino acid sequence NSFGFGG. The PCR reactions were carried out with the sense and anti-sense primers. The reaction was performed under a condition of using a GeneAmp™ PCR kit (Takara Shuzo Co., Ltd.) by adding in 100 µl of reaction solution 20 µM of the primers, respectively, and 1 µg of DNA derived from *A. nidulans*. The reaction program of 35 cycles was performed with each cycle comprising the reaction at 95°C (1 minute), 50°C (1 minute) and 72°C (2 minutes), provided that only in the first cycle, the reaction at 95°C was extended to 3 minutes and the reaction temperature at 50°C was changed into 35°C. After completion of reaction, the reaction mixture was subjected to extraction with 100 µl of chloroform to recovered the aqueous layer and chloroform was then removed with 100 µl of ether to give an aqueous layer, from which 10 µl portion was analyzed by 2% agarose gel electrophoresis.

[0040] As a result, DNA having the same size as the expected one (ca. 330 bp) was detected. The DNA fragment was cloned into a pCRII vector (Invitrogen Co.). The DNA sequence was determined by the dideoxy method with use of a fluorescent sequencer (manufactured by Applied Biosystem Co.). The DNA was labeled with 32P-dCTP using Multiscribe DNA labelling kit (Amersham Co.) to prepare a probe and used for the following experiment of hybridization.

[0041] *E. coli* P2392 was infected with a phage in the DNA library and ca. 10,000 plaques were formed on a plate having a diameter of ca. 15 cm and containing an NZYM medium, which were transferred onto a nylon membrane. Hybridization was carried out by the method described by Sambrook et al. (Molecular Cloning; Second edition, Cold Spring Harbor Laboratory Press, 1989) under a condition in a solution consisting of 5 × SSC (1 × SSC comprising 0.15 M NaCl and 15 mM sodium citrate), 10 mM EDTA, 10 × Denhardt solution (50 × Denhardt solution comprising Ficoll (Type 400, Pharmacia Co.), polyvinyl pyrrolidone, bovine serum albumin (fraction V, Sigma Co.), respectively, in an amount of 10 g/l), and 250 µg/ml of salmon sperm DNA at 60°C for 16 hours. Thereafter, the membrane was washed twice with 5 × SSC and 0.1% SDS solution at 45 °C for 15 minutes and subjected to autoradiography. Ten positive clones were purified to obtain the phage DNAs, which were cleaved with some restriction enzymes, subjected to agarose gel electrophoresis followed by Southern blotting on a nylon membrane according to the conventional method. The membrane was hybridized under the same condition as the above described plaque hybridization to compare the hybridization strengths and the DNA fragment lengths. As a result, the two clones of them (J, B, F) were believed to be satisfactory from the standpoint of both strength and fragment length, and thus cleaved further with some restriction enzymes to carry out Southern hybridization. As a result, a hybridizing fragment having a length of about 5 kbp was detected by cleaving with Sal I, so that it was subcloned to the Sal I site of pUC 19 (Takara Shuzo Co., Ltd.) (referred to as pB and pF, respectively). When each clone was subjected in more detail to mapping with restriction enzymes, pB and pF were judged as the identical DNA. Thus, a deletion plasmid on the latter was prepared with restriction enzyme according to the conventional method, and the base sequence of the DNA chain for about 2 kbp fragment comprising the hybridizing DNA fragment was determined with a fluorescent sequencer (SEQ ID NO. 1). An open reading frame (ORF) consisting of 1251 bp was found from it, and an amino acid sequence having 417 residues was presumed (SEQ ID NO. 2). Comparing the amino acid sequence with that of proteins registered to database for homology, it exhibited significant homology with a fatty acid synthase. In particular, it exhibited the highest homology of 74% with the protein which was thought to be fabF or J in the total genome sequence for *Synechocystis* sp. strain PCC6803 prepared by KAZUSA DNA Institute (database DDBJ accession No. D90905, PID: g1652389) (Fig. 3). As regards the other proteins, it exhibited homologies of 43 - 46% with the KASII gene derived from the castor bean and the KASI gene derived from barley and 35% with KASI derived from *E. coli*. These homologies of genes however could not distinguish which of KASI, II or III corresponds to the function of the ORF.

Example 3 Measurement of activity of KASII-like gene derived from *Anacystis nidulans* in *E. coli*

[0042] In order to specify the function of the above described gene, the expression in *E. coli* was tried. Firstly, in order to remove the excess DNA sequences upstream and downstream ORF, prepared was a DNA in which at the N terminal, *Nde*I site was introduced in the 5' side of the start codon (ATG) by DNA synthesis (Applied Biosystem Co.), while *Hind* III site was introduced immediately downstream the ORF.

55 1: 5'-CGCACATATGACTGAAACCGGACGCC (SEQ ID NO. 5)
2: 5'-CCGCAAGCTTGCAGCAGCGCGTACTGCG (SEQ ID NO. 6)

[0043] PCR reaction was performed with both synthetic DNAs as a primer in the presence of pF as a template DNA.

Regarding the reaction condition, the reaction was repeated 30 cycles according to the manual from Perkin-Elmer Co. with each cycle comprising the reaction at 94°C (1 minute), 60°C (1 minute) and 72°C (2 minutes). The reaction product was cleaved with NdeI and Hind III, followed by cloning with pET17b preliminarily cleaved with the same restriction enzyme set as described above into *E. coli* strain DHS, which was then cloned into a strain BL21(DE3)pLysS (Novagen).

[0044] The recombinant of a latter *E. coli* strain was cultured (32°C) until the turbidity of the culture solution reached 0.5 OD at a wavelength of 600 nm in 75 ml of LB medium to which 100 µg/ml of ampicillin and 30 µg/ml of chloramphenicol had been added. IPTG was then added at a final concentration of 0.4 mM, and culturing was further continued for 2 hours. *E. coli* was recovered from the culture medium by centrifugation at 10,000 × g for 10 minutes, and the cells were washed with 50 mM Tris HCl (pH 7.4) and frozen at -20°C. The cells were thawed in ice with a solution consisting of 20 mM Tris HCl (pH 8.0), 20 mM dithiothreitol, 10 mM MgCl₂ and 1 µg/ml DNase I. The mixture was centrifuged at 100,000 × g at 4°C for 1 hour, and the protein solution as the supernatant was subjected to SDS electrophoresis on a slab gel having a polyacrylamide concentration gradient from 10 to 20% followed by dying with Coomassie brilliant blue. As a result, the protein derived from *Anacystis nidulans* was detected as a protein having a molecular weight of about 50 kDa.

[0045] As for the fatty acid composition of *E. coli*, the fatty acids were recovered for analysis as the methyl esters from the cells cultured as described above. Methylation was carried out by heating about 5 mg of lipid together with 1 ml of 5% hydrochloric acid in anhydrous methanol in a sealed tube in boiling water for 4 hours, the reaction mixture was cooled by standing followed by extracting the fatty acid methyl esters with hexane. Methylated fatty acid esters were analyzed on a capillary column (polyester liquid phase; 10% EGSS-X, 175°C) with a hydrogen flame ionization detector. Fatty acids were determined by comparing their relative retention times with those of standard methylated fatty acids. The results are listed in the following table.

25

30

Fatty acid compositions in <i>E. coli</i>						
Sample	14:0	16:0	16:1	18:0	18:1	16:0+16:1/18:0+18:1
Control	2	37	21	1	38	1.49
Recombinant #1	0	24	13	6	57	0.59
Recombinant #2	0	21	11	7	62	0.46

[0046] As a result, it has been found that the fatty acids with 16 carbon atoms (16:0 and 16:1) were decreased, while the ratio of the fatty acids with 18 carbon atoms (18:0 and 18:1) was increased substantially.

Industrial Applicability

[0047] The DNA chain coding for protein which has a β-ketoacyl-ACP synthetase II enzyme activity represented by the enzyme KASII derived from *Anacystis nidulans* has been provided by the present invention. The gene coding for the enzyme protein of the present invention, as described above, is a gene of the enzyme KASII which has a remarkably high activity of converting a fatty acid (particularly, palmitic acid with C16) into an even longer fatty acid (particularly, stearic acid with C18), and is useful, by using transformation, for the improvement of the lipids of plants, the improvement of the lipids of microorganisms, particularly for the control of the ratio between fatty acids having 16 and 18 carbon atoms or for the increase of the content of unsaturated fatty acids.

50

55

SEQUENCE LISTING

Applicant: Kirin Beer Kabushiki Kaisha

Title of the Invention: β -Ketoacyl-ACP Synthetase II Enzymes
AND Genes Coding for Same

Docket No.: 113702-432

Filing Date: January 20, 1998

Number of Sequences: 6

SEQ ID No.: 1

LENGTH: 1251 base pairs

TYPE: nucleic acid (DNA)

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA

ORIGINAL SOURCE:

ORGANISM: Anacystis nidulans

STRAIN: IAM M-6

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```

ATGACTGAAA CGGGACGCCA GCGTGTGTT ATTACTGGTT TGGGAGCCAT TACTCCCAC 60
GGTAATGATC CAACGGAATA TTGGCAGGGG ATCCCTGGCG GTCGCAACGG CATCGATCTG 120
ATTGGGGCT TTGATGGCTC TCGTCACGCC TGCAAAATTG CCGGGGAGGT CAAGGACTTT 180
GACCCCCACCC AGTACATGGA CCGCAAGGAT GCTAACGGGA TGGATCGGTT TGACACAATG 240
CCGGTTGCTG CCAGTCGCCA AGCAGTCGCC GATGCCAACC TGGACATCAC TGAACGTGAAT 300
GCGGATGCGA TCGGGGTGCT GATCGGCTCA GGCAATTGGT GTTTGAGGGT GATGGAGGAC 360
CAGCAGACGG TTTTGCTGGA AAAAGGCCCC GATCGCTGCA GCCCCCTTCAT GGTGCCGATG 420
ATGATGCCA ACATGGCGGC AGGACTGACG GCCATCCAGT TGGGTGCCAA AGGCCCCTTGC 480
AATGTCACGG TGAATGCTTG CGCTGCCGGT TCTAATGCGG TGGGTGAAGC CTTCCGGCTG 540
ATTCAAGCACG GCTATGCCA AGCCATGATC TGTGGCGGAA CTGAATCCTG TGTGACCCCCA 600
CTGGCTATGG CCGGTTTGC GGCCTGTAAG GCACTGTCGC TGGCAACGA TGACCCGGCC 660
CATGCTTGCC GTCCCTTTGA CCAAGGCCGT GATGGTTTG TGATGGCGA AGGCGCAGGG 720
ATTTTGGTCT TGGAAATCCTT GGACCATGCC CAAGCGAGGG GCGCTCACAT CTATGGCGAA 780
ATCGTCGGCT ATGGCATGAC CTGTGATGCC TATCACATCA CCTCGCCGGT CCCAGGTGGT 840
TTGGGTGCCGG CCCGGGCGAT CGAGTTGGG CTCCGCGATG CCAATCTGCA GCCCAGCCAA 900
GTCAGCTACA TCAATGCTCA CGGCACCAGC ACACCCGCCA ACGACAGGCAC CGAAACGGCA 960
GCTATTAAGA AAGCCCTAGG TGAGCACGCC TACAAAACCG TGATCAGCTC GACTAAGTCG 1020
ATGACCGGTC ACCTGTTAGG GGGCTCCGGC GGAATTGAGG CGGTAGCGGC AACCCCTCGCG 1080
ATCGCTGAGG ACATGGCTCC GCGGACGGATT AACCTGGAAG ATCCCGATCC CGATTCGGAC 1140
TTGGACTATG TCCCCAATCA CGCGCGATCG CTACCGGTGG AAGTGGCTTT GTCCAATTCC 1200

```

TTGGGCTTTC GTGGGCACAA CGTCACGCTG GCCTTCCGGA AATTCCATCC C

1251

SEQ ID No.: 2

LENGTH: 417 amino acids

TYPE: protein

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

ORGANISM: Anacystis nidulans

STRAIN: IAM M-6

NAME/KEY: CDS

IDENTIFICATION METHOD: P

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Glu Thr Gly Arg Gln Arg Val Val Ile Thr Gly Leu Gly Ala

1 S 10 15

Ile Thr Pro Ile Gly Asn Asp Pro Thr Glu Tyr Trp Gln Gly Ile Leu

20 25 30

Ala Gly Arg Asn Gly Ile Asp Leu Ile Arg Gly Phe Asp Ala Ser Arg

35 40 45

His Ala Cys Lys Ile Ala Gly Glu Val Lys Asp Phe Asp Pro Thr Gln

50 55 60

Tyr Met Asp Arg Lys Asp Ala Lys Arg Met Asp Arg Phe Ala Gln Leu

65 70 75 80

Ala Val Ala Ala Ser Arg Gln Ala Val Ala Asp Ala Lys Leu Asp Ile

85 90 95

Thr Glu Leu Asn Ala Asp Ala Ile Gly Val Leu Ile Gly Ser Gly Ile

100 105 110

Gly Gly Leu Arg Val Met Glu Asp Gln Gln Thr Val Leu Leu Glu Lys

115 120 125

Gly Pro Asp Arg Cys Ser Pro Phe Met Val Pro Met Met Ile Ala Asn

130 135 140

Met Ala Ala Gly Leu Thr Ala Ile Gln Leu Gly Ala Lys Gly Pro Cys

145 150 155 160

Asn Val Thr Val Thr Ala Cys Ala Ala Gly Ser Asn Ala Val Gly Glu

165 170 175

Ala Phe Arg Leu Ile Gln His Gly Tyr Ala Gln Ala Met Ile Cys Gly

180 185 190

Gly Thr Glu Ser Cys Val Thr Pro Leu Ala Met Ala Gly Phe Ala Ala
 - 195 200 205
 5 Cys Lys Ala Leu Ser Leu Arg Asn Asp Asp Pro Ala His Ala Cys Arg
 210 215 220
 Pro Phe Asp Gln Gly Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly
 225 230 235 240
 10 Ile Leu Val Leu Glu Ser Leu Glu His Ala Gln Ala Arg Gly Ala His
 245 250 255
 Ile Tyr Gly Glu Ile Val Gly Tyr Gly Met Thr Cys Asp Ala Tyr His
 15 260 265 270
 Ile Thr Ser Pro Val Pro Gly Gly Leu Gly Ala Ala Arg Ala Ile Glu
 275 280 285
 20 Phe Gly Leu Arg Asp Ala Asn Leu Gln Pro Ser Gln Val Ser Tyr Ile
 290 295 300
 Asn Ala His Gly Thr Ser Thr Pro Ala Asn Asp Ser Thr Glu Thr Ala
 305 310 315 320
 25 Ala Ile Lys Lys Ala Leu Gly Glu His Ala Tyr Lys Thr Val Ile Ser
 325 330 335
 Ser Thr Lys Ser Met Thr Gly His Leu Leu Gly Gly Ser Gly Gly Ile
 340 345 350
 30 Glu Ala Val Ala Ala Thr Leu Ala Ile Ala Glu Asp Met Val Pro Pro
 355 360 365
 Thr Ile Asn Leu Glu Asp Pro Asp Pro Asp Cys Asp Leu Asp Tyr Val
 370 375 380
 35 Pro Asn Gln Ala Arg Ser Leu Pro Val Glu Val Ala Leu Ser Asn Ser
 385 390 395 400
 Phe Gly Phe Gly Gly His Asn Val Thr Leu Ala Phe Arg Lys Phe His
 40 405 410 415
 Pro
 417

45 SEQ ID No.: 3

LENGTH: 20 base pairs

TYPE: nucleic acid (DNA)

50 STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

5'-CC(ACGT)CC(AG)AA(ACGT)CC(AG)AA(ACGT)GA(AG)TT-3'

20

SEQ ID No.: 4

LENGTH: 23 base pairs

TYPE: nucleic acid (DNA)

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

5'-GA(AG)GA(AG)GT(ACGT)AA(CT)TA(CT)AT(ACT)AA(CT)GC-3'

23

SEQ ID No.: 5

LENGTH: 26 base pairs

TYPE: nucleic acid (DNA)

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION: SEQ ID NO: 5:

5' -CGCACATATGACTGAAACCGGACGCC-3'

26

SEQ ID No.: 6

LENGTH: 27 base pairs

TYPE: nucleic acid (DNA)

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid (synthetic DNA)

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

5' -CCGCAAGCTTGCAGCAGCGCGTACTGC-3'

27

50

Claims

- 55 1. A protein which has an amino acid sequence represented by SEQ ID NO. 2 or substantially the same amino acid sequence as the one represented by SEQ ID NO. 2 and has the β -ketoacyl-ACP synthetase II enzyme activity.
2. A protein according to Claim 1, wherein substantially the same amino acid sequence is the one in which one or

more amino acids are substituted, deleted, inserted or added.

3. A protein according to Claims 1 or 2, wherein the amino acid sequence is derived from cyanobacterium.
4. A gene coding for a protein which has an amino acid sequence represented by SEQ ID NO. 2 or substantially the same amino acid sequence as the one represented by SEQ ID NO. 2 and has the β -ketoacyl-ACP synthetase II enzyme activity.
- 10 5. A β -ketoacyl-ACP synthetase II enzyme active protein gene according to claim 4, wherein substantially the same amino acid sequence is the one in which one or more amino acids are substituted, deleted, inserted or added.
6. A gene according to Claims 4 or 5, wherein the protein encoded is derived from cyanobacterium.
- 15 7. A recombinant vector comprising the β -ketoacyl-ACP synthetase II enzyme active protein gene according to any one of Claims 4 -6.
8. A cell in which the β -ketoacyl-ACP synthetase II enzyme active protein gene according to any one of Claims 4 -6 has been introduced.

20

25

30

35

40

45

50

55

EP 0 969 014 A1

M TETGRQRWV ITGLGAI TP I GNDPTEYHOG I LAGRNGIDL I RGFDASRHA CKIAGEVKDF
DPTQYMDRKD AKRMDRFAQL AVAASRQAVA DAKLDITELN ADAIGVLIGS GIGGLRVMED
QQTVLLEKGP DRCSPFMVPM MIANMAAGLT A1QLGAKGPC NVTVTACAAG SNAVGEAFRL
IQHGYAQAMI CGGTESCVP TP LAMAGFAACK ALSLRNDDPA HACRPFDOGR DGFVMGEGAG
ILVLESLEHA QARGAHIFYGE I VGYGMTCDA YHITSPVPGG LGAARAIEFG LRDANLOPSQ
VSYINAHGTS TPANDSTETA AIKKALGEHA YKTVISSTKS MTGHLLGGSG GIEAVAATLA
IAEDMVPPTI NLEDPDPCD LDYVPNOARS LPVEVALSNS FGFGGHNVTL AFRKFHP

FIG. 1

	10	20	30	40	50		
BARLEY	MHAAAHALGLRVPPPAPRRRARPRR—PAAAVLATSAAPORE—TDP—RKRVV						
CASTOR-OIL PLANT	PCSHYYSSNGLFPNTPLLPKRHPRLHHRLPRSGEAAVAVOPEKEVATNKKPLUKORVV	80	90	100	110	120	130
BARLEY	I TGWGLASVFGSOVDTFYDRLLAGESGVGPIDRFDASSFPTRFAGOIRGFSEGYIDGKN	60	70	80	90	100	110
CASTOR-OIL PLANT	VTGMGVVSPLGHIDVYNNL LDGSSG ISQIDSFDCAQFPTRIAGEIKSFSTDGIVAPKL	140	150	160	170	180	190
BARLEY	DRRLDOCIRYCILSGKKALESAGLGAGSDAHVKLDVGRAGVLVGTGMGGLSVSDGVQNL	120	130	140	150	160	170
CASTOR-OIL PLANT	SKRMOKFMLYMLTAGKKALADGGI—TEDHMDELDKARCGVLIGSAMGGMKVFNDIAEAL	200	210	220	230	240	
BARLEY	IEKGYRKISPFPIPYAINTNGSALLAIDVGFMGPNSISTACATSNYCFYAANHTRGE	180	190	200	210	220	230
CASTOR-OIL PLANT	-RISYRKWMNPFCVPFATTNNGSAMLAQOLGMGPNSISTACATSNFCILNAANHTRGE	250	260	270	280	290	300
BARLEY	ADIIIVAGGTEAAIIIPIGLGGFVACRALSORNDOPITACRPWDKERDGFMGEGAGVLVME	240	250	260	270	280	290
CASTOR-OIL PLANT	ADIMLCGGSDAAIIIPIGLGGFVACRALSORNDPTKASRPWDNRDGFVMGEGAGVLLLE	310	320	330	340	350	360
BARLEY	SLEHAMKRDAPIIAEYLGGAVNCDAYHMTDPRADGLGVSSCTMSLRDAGVAPEEVNYIN	300	310	320	330	340	350
CASTOR-OIL PLANT	ELEHAKKRGANIYAEFLGGSFTCDAYHMTEPRPDGVGVLICIEKALARSGVSKEEVNYIN	370	380	390	400	410	420
BARLEY	AHATSTLAGOLAEVRAIKOVFKNPSEIKINSTKSMIGHCLGAAGGLEIATIKSITTGWV	360	370	380	390	400	410
CASTOR-OIL PLANT	AHATSTPAGOLKEYALIACFSQNPDLRVNSTKSMIGHLLGAAGAVEIATIQAIRTGWV	430	440	450	460	470	480
BARLEY	HPTINQFNPEPEVDFDTVANEKKOH-EVNNGISNSFGGGHNSWWFAPFK	420	430	440	450	460	
CASTOR-OIL PLANT	HPNINLENPEEGVDTKVLYGPKKERLDIKVALNSNSFGGGHNSSIIFAPYK	490	500	510	520	530	

FIG. 2

FIG. 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP98/00194

A. CLASSIFICATION OF SUBJECT MATTER Int.Cl' C07K14/405, C12N9/10, C12N15/54, C12N15/31, C12N1/21 // C12N1/12, (C12N15/54, C12R1:89), (C12N9/10, C12R1:19) <small>According to International Patent Classification (IPC) or to both national classification and IPC</small>		
B. FIELDS SEARCHED <small>Minimum documentation searched (classification system followed by classification symbols)</small> Int.Cl' C07K14/405, C12N9/10, C12N15/54, C12N15/31, C12N1/21, C12N1/12		
<small>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</small>		
<small>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)</small> BIOSIS (DIALOG), WPI (DIALOG), GenBank/EMBL (GENETYX), SwissProt (GENETYX)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP, 7-501446, A (E.I. Du Pont de Nemours & Co.), February 16, 1993 (16. 02. 93) & WO, 93/10240, A1 & AU, 9230735, A & EP, 667906, A1 & US, 5500361, A	1-8
X	JP, 6-500234, A (Calgene Inc.), January 13, 1994 (13. 01. 94) & WO, 92/03564, A1 & EP, 495096, A1 & US, 5475099, A	1-8
A	WORSHAM, L.M.S. et al., "Early Catalytic Steps Of Euglena-Gracilis Chloroplast Type II Fatty Acid Synthase", BIOCHEM. BIOPHYS. ACTA (1993) Vol. 1170, No. 1, p.62-71	1-8
A	GARWIN, J.L. et al., "Structural, Enzymatic, and Genetic Studies of β -Ketoacyl-Acyl Carrier Protein Synthases I and II of Escherichia coli", J. Biol. Chem (1980) Vol. 255, No. 24, p.11949-11956	1-8
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document not published on or after the international filing date which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "L" document referring to an oral disclosure, i.e., exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed <small>T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</small> <small>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</small> <small>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</small> <small>"A" document member of the same patent family</small>		
Date of the actual completion of the international search April 15, 1998 (15. 04. 98)	Date of mailing of the international search report April 28, 1998 (28. 04. 98)	
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer	
Faxsimile No.	Telephone No.	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP98/00194

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MAGNUSON, K. et al., "The Putative fabJ Gene of Escherichia coli Fatty Acid Synthesis Is the fabF Gene", J. BACTERIOL, (1995) Vol. 177, No. 12, p.3593-3595	1-8
A	SIGGAARD-ANDERSEN, M. et al., "The fabJ-encoded β -ketoacyl-[acyl carrier protein] synthase IV from Escherichia coli is sensitive to cerulenin and specific for short-chain substrates", Proc. Natl. Acad. Sci. USA, (1994) Vol. 91, No. 23, p.11027-11031	1-8
A	SHEN, Z. et al., "Isolation of Vibrio harveyi Acyl Carrier Protein and the fabG, acpP, and fabF Genes Involved in Fatty Acid Biosynthesis", J. BACTERIOL, (January 1996) Vol. 178, No. 2, p.571-573	1-8

